Claims

What is claimed is:

5   1.   A method for segmenting a music video (507) in a multimedia stream (505), said method comprising:

receiving a multimedia stream (505) including at least one music video (507);

segmenting said at least one music video (507) from said multimedia stream (505) by evaluating a plurality of content features (1210, 1220, 1230) related to said
10   multimedia stream (505); and

identifying said at least one music video (507).

2.   The method of claim 1, further comprising the step of generating a summary (410) of said at least one music video (507).

3.   The method of claim 1, wherein said summary (410) of said at least one
15   music video (507) is presented to a user based on personalized preferences.

4.   The method of claim 1, wherein said at least one music video (507) may be retrieved by a user based on personalized preferences.

5.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) are processed using a pattern recognition engine (1000) to identify said at least
20   one music video (507).

6.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) are processed using a Bayesian Belief Network (1000) to identify said at least one music video (507).

7.   The method of claim 1, wherein said plurality of content features (1210,
25   1220, 1230) are processed using one or more video segmentation rules (1115) to identify said at least one music video (507).

8.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) includes a face presence feature to evaluate patterns in the presentation of faces in said multimedia stream (505).

5   9.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) includes a videotext presence feature that determines when videotext appears in said multimedia stream (505).

10.   The method of claim 1, wherein said plurality of content features (1210, 10   1220, 1230) includes a color histogram feature to evaluate patterns in the color content of said multimedia stream (505).

11.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) includes a camera cut feature to evaluate patterns in the camera cuts and 15   movements in said multimedia stream (505).

12.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) includes an analysis of key words obtained from a transcript of said at least one music video (507).

20

13.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) includes an analysis of low level features derived directly from said multimedia stream.

25   14.   The method of claim 13, wherein said low level features include one or more of a number of edges or shapes or local or global motion.

15.   The method of claim 1, wherein said plurality of content features (1210, 1220, 1230) includes an audio feature.

30

16.   The method of claim 15, wherein said audio feature evaluates a volume of said multimedia stream (505).

17.   The method of claim 15, wherein said audio feature evaluates one or more of a mel frequency cepstral coefficient (MFCC), linear predictive coefficient (LPC), or variations in pitch bandwidth, volume or tone.

18.   The method of claim 1, further comprising the step of obtaining identifying information for said at least one music video from an external source.

19.   A method for detecting a chorus in at least one music video (507), comprising the steps of:

          receiving a multimedia stream (505) including said at least one music video (507);

          accessing a transcript associated with said at least one music video (507); and

          detecting said chorus based upon a repetition of words in said transcript.

20.   The method of claim 19, wherein said transcript is obtained from closed caption information.

21.   The method of claim 19, wherein said chorus is employed for an automatic generation of a summary (410) of said at least one music video (507).

22.   The method of claim 19, further comprising the steps of detecting and clustering said repeated words.

23.   The method of claim 19, wherein said detecting step is further based upon additional content features related to said multimedia stream.

24.   The method of claim 19, further comprising the step of obtaining identifying information for said at least one music video from an external source.

25.   An apparatus for segmenting a music video (507) in a multimedia stream (505), said apparatus comprising:

          a memory (280); and

at least one controller (270), coupled to the memory (280), operative to:

receive a multimedia stream (505) including at least one music video (507);

apply a plurality of content features (1210, 1220, 1230) related to said multimedia stream (505) to a pattern recognition engine (1000) to segment said at least one

5   music video (507) from said multimedia stream (505); and

identify said at least one music video (507).

26.   The apparatus of claim 25, wherein said pattern recognition engine (1000) is a Bayesian Belief Network.

27.   The apparatus of claim 25, wherein said pattern recognition engine (1000) is

10   a neural network.

28.   The apparatus of claim 25, wherein said pattern recognition engine (1000) employs an Auto Regressive Moving Average technique.

29.   The apparatus of claim 25, wherein said plurality of content features (1210, 1220, 1230) includes at least two of a face presence feature; a videotext presence feature; a

15   color histogram feature; a camera cut feature; and an analysis of key words obtained from a transcript of said at least one music video (507).

30.   An apparatus for segmenting a music video (507) in a multimedia stream (505), said apparatus comprising:

20   a memory (280); and

at least one controller (270), coupled to the memory (280), operative to:

.receive a multimedia stream (505) including at least one music video (507);

apply a plurality of content features (1210, 1220, 1230) related to said multimedia stream (505) to one or more video segmentation rules (1115) to segment said at

25   least one music video (507) from said multimedia stream (505); and

identify said at least one music video (507).

31.   The apparatus of claim 30, wherein said plurality of content features (1210, 1220, 1230) includes at least two of a face presence feature; a videotext presence feature; a color histogram feature; a camera cut feature; and an analysis of key words obtained from a transcript of said at least one music video (507).

5

32.   The apparatus of claim 30, wherein said one or more video segmentation rules (1115) define a threshold for said plurality of content features (1210, 1220, 1230) to determine when a video segment has occurred.

10